

Portfolio Choice with Transaction Costs: a User's Guide *

Paolo Guasoni[†]

Johannes Muhle-Karbe[‡]

August 1, 2012

Abstract

Recent progress in portfolio choice has made a wide class of problems involving transaction costs tractable. We review the basic approach to these problems, and outline some directions for future research.

Mathematics Subject Classification: (2010) 91G10, 91G80.

JEL Classification: G11, G12.

Keywords: transaction costs, long-run, portfolio choice, Merton problem.

*The authors are grateful to Ren Liu for proofreading the manuscript.

[†]Boston University, Department of Mathematics and Statistics, 111 Cummington Street Boston, MA 02215, USA.
 Dublin City University, School of Mathematical Sciences, Glasnevin, Dublin 9, Ireland, email: guasoni@bu.edu.
 Partially supported by the ERC (278295), NSF (DMS-0807994, DMS-1109047), SFI (07/MI/008, 07/SK/M1189, 08/SRC/FMC1389), and FP7 (RG-248896).

[‡]ETH Zürich, Departement Mathematik, Rämistrasse 101, CH-8092, Zürich, Switzerland, and Swiss Finance Institute, email: johannes.muhle-karbe@math.ethz.ch. Partially supported by the National Centre of Competence in Research “Financial Valuation and Risk Management” (NCCR FINRISK), Project D1 (Mathematical Methods in Financial Risk Management), of the Swiss National Science Foundation (SNF).

1 Introduction

Transaction costs, originally considered one of many imperfections that are best neglected, have now become a very active and fast-growing theme in Mathematical Finance.¹ From the outset, such a growth may seem puzzling, since over the same period transaction costs have dramatically declined across financial markets, as stock exchanges have been fully automated, and paper trades replaced by electronic settlements. In fact, the interest for transaction costs reflects both the increased attention to robustness of financial models, and the growing role of high-frequency trading. The decline of bid-ask spreads has sparked a huge increase in trading volume, and high-volume strategies require a careful understanding of the effects of frictions on their returns.

At the same time, transaction costs help understand trading volume itself. In frictionless models, investors continuously rebalance their portfolios, as to hold a constant mix of assets over time. Since trading volume is proportional to the total variation of a portfolio, and prices follow diffusions that have infinite variation, such models lead to the absurd conclusion that trading volume is infinite over any time interval. With transaction costs, even small trading costs make it optimal for investors to trade infrequently, allowing wide oscillations in their portfolios.

This paper reviews a recent approach, which has made portfolio choice with transaction costs more tractable, and which appears to be applicable in more complex settings. This approach is not based on any new revolutionary concept, but it rather tries to combine several ideas that were previously used in isolation. Thus, we present a new toolbox that contains several used tools. In a nutshell, we argue that a natural approach to portfolio choice problems with transaction costs entails four steps: (i) heuristic control arguments to identify the long-run value function, (ii) construction of a candidate shadow price using marginal rates of substitution, (iii) verification and finite-horizon bounds using the myopic probability, and (iv) asymptotic results from the implicit function theorem.

The advantages of this approach are threefold. First, it combines the dimension-reduction and higher tractability of the long-horizon problem with exact finite-horizon bounds, which keep a firm grip on the robustness of the solution. Second, we show that the free-boundaries arising with transaction costs can sometimes be identified explicitly in terms of a single parameter, the *equivalent safe rate*, which remains the only non-explicit part of the solution. This reduction is useful both for theoretical and for practical purposes, as it helps to simplify proofs as well as calculations. Third, this approach leads to the simultaneous computation of several related quantities, such as welfare, portfolios, liquidity premia, and trading volume.

The paper proceeds as follows: in the next section, we present a brief timeline of related research, which is far from exhaustive, and only aims at putting the paper in context. The following section introduces the main problem, discussing the relative advantages of the three main models with terminal wealth, consumption, and long-horizon. This section also discusses the typical heuristic arguments of stochastic control that lead to an educated guess for the value function, and the identification of the corresponding free boundaries. For the long-run problem, the following section shows the passage from the heuristic calculations to a verification, which relies on two central ideas. The first one is the construction of a *shadow* market, an imaginary frictionless market, built to deliver the same optimal strategy as the original market with transaction cost. This shadow market harnesses transaction costs by hiding them inside a more complex model, without transaction costs, but in which investment opportunities are driven by a state variable that represents the portfolio composition of the investor. This insight – that transaction costs are essentially equivalent to state-dependent investment opportunities – in turn allows to exploit the approach to verification based

¹The Mathscinet database shows only nine publications with “transaction costs” in their title in the eighties (1980-1989). This figure rises to 52 in the nineties (1990-1999), and to 278 in the naughties (2000-2009).

on the change of measure to the myopic probability.

We conclude with a deliberately speculative section on three open problems: multiple assets, return predictability, and option spreads. We argue that with transaction costs, multivariate models present both a substantial technical challenge, and a potentially fertile ground for novel financial insights, which may alter the conventional wisdom on fund separation. Likewise, transaction costs may help reconcile statistical evidence on return predictability with the poor out-of-sample performance of market-timing strategies. Finally, the large bid-ask spreads observed in options on highly liquid assets still lack a theoretical basis, and transaction costs are a natural avenue to search for an explanation.

2 Literature Review

Portfolio choice with transaction cost starts with the seminal papers of Magill and Constantinides (1976), Constantinides (1986), and Dumas and Luciano (1991), in the wake of the frictionless results of Merton (1969; 1971). From heuristic arguments, these early studies gleaned central insights that held up to subsequent formal proofs. First, optimal portfolios entail a no-trade region, in which it is optimal to keep existing holdings in all assets. Optimal portfolios always remain within this region, and hence trading should merely take place at its boundaries. The no-trade region is wide, even for small transaction costs, implying that investors should accept wide fluctuations around the frictionless target. Second, the large no-trade region has a small welfare impact (Constantinides, 1986), because the displacement loss is small near the frictionless optimum, and the wide no-trade region minimizes the effect of transaction costs.

On the mathematical side, Taksar, Klass, and Assaf (1988) reduce the maximization of logarithmic utility from terminal wealth at a long horizon to the solution of a nonlinear second-order ODE with free boundaries, to be determined numerically. Davis and Norman (1990) accomplish this feat for power utility from consumption with infinite horizon. Shreve and Soner (1994) extend their analysis with viscosity techniques, removing some parametric restrictions. Shreve and Soner (1994) and Rogers (2004) study the size of the no-trade interval and the utility loss due to transaction costs. They argue that these are of order $O(\varepsilon^{1/3})$ and $O(\varepsilon^{2/3})$, respectively, where ε is the proportional cost, in line with the numerical results of Constantinides (1986) alluded to above. Building on earlier heuristic results of Whalley and Wilmott (1997), Janeček and Shreve (2004) explicitly determine the coefficients of the leading-order corrections around the frictionless case $\varepsilon = 0$. In a general Markovian setting and for arbitrary utility functions, Soner and Touzi (2012) characterize the corresponding quantities in terms of an ergodic control problem. All these papers employ stochastic control as their main tool.

A new strand of literature, which finds its roots in the seminal work of Jouini and Kallal (1995) and Cvitanić and Karatzas (1996), seeks to bring martingale methods, now well-understood in frictionless markets, to bear on transaction costs. This idea has already shown its promise in the context of superreplication: Guasoni, Rásonyi, and Schachermayer (2008) prove the face-lifting theorem of Soner, Shreve, and Cvitanić (1995) for general continuous processes, using an argument based on shadow prices. Kallsen and Muhle-Karbe (2010) explore this approach for optimal consumption from logarithmic utility, showing how shadow prices simplify verification theorems. Gerhold, Muhle-Karbe, and Schachermayer (2011b) exploit this idea to obtain the expansions of Janeček and Shreve (2004) for logarithmic utility, but with an arbitrary number of terms. The present study reviews a streamlined version of the approach put forward by Gerhold et al. (2011a), who prove a verification theorem and derive full asymptotics for the optimal policy, welfare, and implied trading volume in the long-run model of Dumas and Luciano (1991). The duality-based

verification is based on applying the frictionless long-run machinery of Guasoni and Robertson (2012) to a fictitious shadow price, traded without transaction costs. Compared to Gerhold et al. (2011a), finding a candidate shadow price is greatly simplified by applying an observation originally made by Loewenstein (2000): Given a smooth candidate value function, it can simply be obtained via the marginal rate of substitution of risky for safe assets for the frictional investor. (Also cf. Herzegh and Prokaj (2011); Liu and Muhle-Karbe (2012) for applications of this idea to related problems.)

3 The Basic Model

3.1 Objectives

Let X_t^π denote the wealth of an investor who follows the portfolio π_t , and let c_t his consumption rate, both at time t . The three typical objectives for portfolio choice with power utility $U(x) = x^{1-\gamma}/(1-\gamma)$ are:

$$\max_{\pi} E \left[\frac{(X_T^\pi)^{1-\gamma}}{1-\gamma} \right], \quad (\text{terminal wealth}) \quad (3.1)$$

$$\max_{\pi, c} E \left[\int_0^\infty e^{-\delta t} \frac{c_t^{1-\gamma}}{1-\gamma} dt \right], \quad (\text{consumption}) \quad (3.2)$$

$$\max_{\pi} \liminf_{T \rightarrow \infty} \frac{1}{T} \log E [(X_T^\pi)^{1-\gamma}]^{\frac{1}{1-\gamma}}. \quad (\text{long run}) \quad (3.3)$$

Expected utility from terminal wealth (3.1) has attracted the attention of most of the semi-martingale literature (see, for example, Kramkov and Schachermayer (1999) and the references therein). This objective is the simplest for abstract questions, such as existence, uniqueness, well-posedness, and stability, which are by now largely understood. It is also relevant for problems such as retirement planning, which entail a known horizon and no intermediate consumption. Expected utility from intertemporal consumption (3.2) is more appealing for applications to macroeconomics, because it yields an endogenous consumption process c_t , and therefore has testable implications for consumption data.

The long run objective is probably the least intuitive, in view of the limit in (3.3). To understand its economic interpretation, note first that, for a fixed horizon T , the quantity $E [(X_T^\pi)^{1-\gamma}]^{\frac{1}{1-\gamma}}$ coincides with the *certainty equivalent* $U^{-1}(E[U(X_T^\pi)])$ of the payoff X_T^π . If we match this certainty equivalent with $xe^{\rho_T T}$, that is, the investor's initial capital x compounded at some constant rate ρ_T for the same horizon T , we recognize that:

$$\rho_T = \frac{1}{T} \log E [(X_T^\pi)^{1-\gamma}]^{\frac{1}{1-\gamma}}.$$

Thus, the limit in (3.3) has the interpretation of an *equivalent safe rate*, that is, the hypothetical safe rate that would make the investor indifferent between investing optimally in the market, and leaving all wealth invested at this hypothetical rate.

Both the consumption and long-run problems are *stationary* objectives, in that they lead to time-independent solutions (as long as investment opportunities are also stationary). Of course, the advantage of stationary problems is that the resulting optimization problems have one less dimension than similar nonstationary problems, such as utility maximization from terminal wealth. Both objectives model an investor with an infinite horizon, but with some important differences. First, the consumption objective involves the additional time-preference rate δ , which does not

appear in the long-run objective. Second, in typical models (even in a Black-Scholes market, compare Choi, Sirbu, and Žitković (2012)), the consumption objective may not be well posed if risk aversion γ is less than the logarithmic value of one, and investment opportunities are sufficiently attractive. By contrast, the long-run objective is typically well-posed under more general conditions.

The irony of portfolio choice is that its most natural objectives are also the least tractable: the terminal-wealth problem admits closed-form solutions only in rare cases (cf., e.g., Liu (2007)). Even when such solutions exist, they are often too clumsy to yield clear insights on the role of preference and market parameters. Unfortunately, the consumption objective admits explicit solutions primarily in complete markets, or with investment opportunities independent of asset prices, a fact that severely limits our understanding of the effects of partial return predictability on consumption.

The good news is that the long-run problem admits explicit solutions in many situations in which the other two problems do not, its optimal portfolio is almost optimal even for the other objectives, and bounds on the resulting utility loss are available. This general insight is crucial in markets with frictions, such as transaction costs.

3.2 Control Heuristics

We now examine the differences in the Hamilton-Jacobi-Bellman equations arising from the three objectives (3.1), (3.2), and (3.3), in the basic model with one safe asset growing at the riskless rate $r \geq 0$, and a risky asset with ask (buying) price S_t following geometric Brownian Motion:

$$\frac{dS_t}{S_t} = (\mu + r)dt + \sigma dW_t, \quad \mu, \sigma > 0. \quad (3.4)$$

The bid (selling) price is $(1 - \varepsilon)S_t$, where $\varepsilon \in (0, 1)$ is the relative bid-ask spread.

Denote the number of units of the safe asset by φ_t^0 and write the number of units of the risky asset $\varphi_t = \varphi_t^\uparrow - \varphi_t^\downarrow$ as the difference between cumulative purchases and sales. The values of the safe position X_t and of the risky position Y_t (quoted at the ask price) evolve as:

$$dX_t = rX_t dt - S_t d\varphi_t^\uparrow + (1 - \varepsilon)S_t d\varphi_t^\downarrow, \quad (3.5)$$

$$dY_t = (\mu + r)Y_t dt + \sigma Y_t dW_t + S_t d\varphi_t^\uparrow - S_t d\varphi_t^\downarrow. \quad (3.6)$$

The second equation prescribes that risky wealth earns the return on the risky asset, plus units purchased, and minus units sold. In the first equation the safe position earns the safe rate, minus the units used for purchases (at the ask price S_t), and plus the units used for sales (at the bid price $(1 - \varepsilon)S_t$).

For the maximization of utility from terminal wealth, denote the value function as $V(t, x, y)$, which depends on time t , on the safe position x , and on the risky position y . Itô's formula yields:

$$dV(t, X_t, Y_t) = V_t dt + V_x dX_t + V_y dY_t + \frac{1}{2} V_{yy} d\langle Y, Y \rangle_t \quad (3.7)$$

$$= \left(V_t + rX_t V_x + (\mu + r)Y_t V_y + \frac{\sigma^2}{2} Y_t^2 V_{yy} \right) dt \quad (3.8)$$

$$+ S_t(V_y - V_x) d\varphi_t^\uparrow + S_t((1 - \varepsilon)V_x - V_y) d\varphi_t^\downarrow + \sigma X_t V_y dW_t, \quad (3.9)$$

By the martingale optimality principle of stochastic control, the value function $V(t, X_t, Y_t)$ must be a supermartingale for any choice of purchases and sales $\varphi_t^\uparrow, \varphi_t^\downarrow$. Since these are increasing processes, this implies $V_y - V_x \leq 0$ and $(1 - \varepsilon)V_x - V_y \leq 0$, which means that

$$1 \leq \frac{V_x}{V_y} \leq \frac{1}{1 - \varepsilon}. \quad (3.10)$$

In the interior of this “no-trade region”, where the number $\varphi_t = \varphi_t^\uparrow - \varphi_t^\downarrow$ of risky shares remains constant, the drift of $V(t, X_t, Y_t)$ cannot be positive, and must become zero for the optimal policy. This leads to the HJB equation:

$$V_t + rX_tV_x + (\mu + r)Y_tV_y + \frac{\sigma^2}{2}Y_t^2V_{yy} = 0 \quad \text{if} \quad 1 < \frac{V_x}{V_y} < \frac{1}{1-\varepsilon}. \quad (3.11)$$

Next, the value function is homogeneous in wealth, i.e. $V(t, X_t, Y_t) = (X_t)^{1-\gamma}v(t, Y_t/X_t)$, whence setting $z = y/x$:

$$\frac{\sigma^2}{2}z^2v_{zz} + \mu zv_z + r(1-\gamma)v + v_t = 0 \quad \text{if} \quad 1+z < \frac{(1-\gamma)v(t,z)}{v_z(t,z)} < \frac{1}{1-\varepsilon} + z. \quad (3.12)$$

Now, suppose that the no-trade region $\{(t, z) : 1+z \leq \frac{(1-\gamma)v(t,z)}{v_z(t,z)} \leq \frac{1}{1-\varepsilon} + z\}$ coincides with some interval $l(t) \leq z \leq u(t)$ to be found. At $l(t)$ the left inequality in (3.12) holds as equality, while at $u(t)$ the right inequality holds as equality, leading to the boundary conditions:

$$(1+l)v_z(t, l) - (1-\gamma)v(t, l) = 0, \quad (3.13)$$

$$(1/(1-\varepsilon) + u)v_z(t, u) - (1-\gamma)v(t, u) = 0. \quad (3.14)$$

These conditions are not sufficient to identify the solution to the optimization problem, since they can be matched for any trading boundary $l(t), u(t)$. The optimal boundaries are identified as the ones that satisfy the smooth-pasting conditions. These conditions can be seen as limits of the optimality conditions for an impulse control problem with a infinitesimally small cost (Dumas (1991)). In practice, they are derived by differentiating (3.13) and (3.14) with respect to z at the respective boundaries $z = l$ and $z = u$:

$$(1+l)v_{zz}(t, l) + \gamma v_z(t, l) = 0, \quad (3.15)$$

$$(1/(1-\varepsilon) + u)v_{zz}(t, u) + \gamma v_z(t, u) = 0. \quad (3.16)$$

This system defines a *two-dimensional, linear* free-boundary problem in $(t, z) \in [0, T] \times \mathbb{R}$, which is not tractable in general. Liu and Loewenstein (2002) obtain a semiexplicit solution with the randomization approach used by Carr (1998) to price American options.

With utility maximization from infinite-horizon consumption, the value function depends only on the safe and risky positions X_t, Y_t – the problem is stationary. Calculations are similar, with some minor differences: first, the self-financing condition (3.6) must include the term $-c_t dt$ in the cash balance, to account for consumption expenditures. Then, the martingale optimality principle takes the following slightly different form. Since utility is not only incurred at maturity but from consumption along the way, not the value function itself but the sum of past consumption $\int_0^t e^{-\delta u} c_u^{1-\gamma}/(1-\gamma) du$ and the value function $e^{-\delta t} V(X_t, Y_t)$, representing future consumption, should be a supermartingale for any policy and a martingale for the optimizer. Here, δ is the time-preference parameter in (3.2). Then, the term V_t in (3.8) and in turn (3.11) has to be replaced by $\frac{c_t^{1-\gamma}}{1-\gamma} - c_t V_x - \delta V$. Pointwise maximization yields the optimal consumption rate $c_t = V_x^{-1/\gamma}$. Plugging this expression back into the HJB equation and accounting for homogeneity in wealth, the corresponding free-boundary problem for the reduced value function $v(z)$ then reads as:

$$\frac{\sigma^2}{2}z^2v_{zz} + \mu zv_z + ((1-\gamma)r - \delta)v + \frac{\gamma}{1-\gamma}((1-\gamma)v - zv_z)^{1-\frac{1}{\gamma}} = 0 \quad \text{if} \quad 1+z < \frac{(1-\gamma)v(z)}{v_z(z)} < \frac{1}{1-\varepsilon} + z. \quad (3.17)$$

This is the *one-dimensional, nonlinear* free-boundary problem studied by Davis and Norman (1990), who prove a verification theorem, and find a numerical solution. Still, this problem is nontrivial, because the free boundaries points l, u are not easy to identify in terms of the model parameters, and the second order, nonlinear equation (3.17) does not admit a known explicit solution for given initial conditions.

The long-run problem (3.3) gives the best of both worlds, and more. But it requires more audacious heuristics, and nonstandard arguments to be made precise. Puzzlingly enough, this approach was proposed very early in the transaction costs literature by Taksar, Klass, and Assaf (1988) and Dumas and Luciano (1991), but its potential has not become clear until recently.

We start from equation (3.11), derived for a fixed horizon T , and note that the value function V should grow exponentially with the horizon. This observation, combined with homogeneity in wealth, leads to guess a solution of the form $V(t, X_t, Y_t) = (X_t)^{1-\gamma} v(Y_t/X_t) e^{-(1-\gamma)(r+\beta)t}$. It is clear that such a guess in general does not solve the finite-horizon problem, as it fails to satisfy its terminal condition. But it is reasonable to expect that it governs the long-run problem, for which the horizon never approaches. With the above guess, the HJB equation reduces to

$$\frac{\sigma^2}{2} z^2 v''(z) + \mu z v'(z) - (1-\gamma)\beta v(z) = 0 \quad \text{if} \quad 1+z < \frac{(1-\gamma)v(z)}{v'(z)} < \frac{1}{1-\varepsilon} + z, \quad (3.18)$$

which is a *one-dimensional, linear* free-boundary problem – the best of both worlds. Note that in this system β is not exogenous, but an unknown parameter that determines the growth rate of the value function, and which has to be found along with the free boundaries l, u . Indeed, $r + \beta$ is the *equivalent safe rate* which makes a long-term investor indifferent between the original market and this alternative rate alone.

The two crucial advantages of the long-run problem are that the free boundaries l, u have explicit formulas in terms of β , and that it reduces to solving a Cauchy problem for a *first-order* ordinary differential equation. Depending on the problem at hand, this equation may even have an explicit solution, a fact that is useful although not essential for asymptotics. To derive the free boundary l , substitute first (3.15) and then (3.13) into (3.18) to obtain

$$-\frac{\sigma^2}{2}(1-\gamma)\gamma \frac{l^2}{(1+l)^2} v + \mu(1-\gamma) \frac{l}{1+l} v - (1-\gamma)\beta v = 0.$$

Now, observe that $\pi_- = l/(1+l)$ is precisely the risky portfolio weight at the buy boundary, evaluated at the ask price. Factoring out $(1-\gamma)v$, it follows that (cf. Dumas and Luciano (1991))

$$-\frac{\gamma\sigma^2}{2}\pi_-^2 + \mu\pi_- - \beta = 0. \quad (3.19)$$

Likewise, a similar calculation for u shows that the other root of (3.19) is $\pi_+ = u(1-\varepsilon)/(1+u(1-\varepsilon))$, which coincides with the risky portfolio weight at the sell boundary, evaluated at the bid price.

After these calculations, the boundaries l, u , or equivalently π_-, π_+ , are uniquely identified as solutions of the above equation, once the parameter β is found:

$$\pi_{\pm} = \frac{\mu \pm \sqrt{\mu^2 - 2\gamma\sigma^2\beta}}{\gamma\sigma^2}.$$

The formulas become even clearer by replacing the parameter β with $\lambda = \sqrt{\mu^2 - 2\gamma\sigma^2\beta}$. With this notation, in which $\lambda = 0$ corresponds to the frictionless setting, $\beta = (\mu^2 - \lambda^2)/2\gamma\sigma^2$ and the buy and sell boundaries have the intuitive representation

$$\pi_{\pm} = \frac{\mu \pm \lambda}{\gamma\sigma^2}, \quad (3.20)$$

from which $l = \frac{\pi_-}{1-\pi_-}$ and $u = \frac{1}{1-\varepsilon} \frac{\pi_+}{1-\pi_+}$ are obtained directly. Thus, it remains to find λ to identify both the free-boundaries and the equivalent safe rate $r + \beta$. To this end, it is convenient to apply the substitution

$$v(z) = e^{(1-\gamma) \int_0^{\log(z/l(\lambda))} w(y) dy}, \quad \text{i.e.,} \quad w(y) = \frac{l(\lambda) e^y v'(l(\lambda) e^y)}{(1-\gamma) v(l(\lambda) e^y)},$$

which reduces the free-boundary problem to a Cauchy problem with a terminal condition:

$$w'(y) + (1-\gamma)w(y)^2 + \left(\frac{2\mu}{\sigma^2} - 1\right) w(y) - \gamma \left(\frac{\mu - \lambda}{\gamma\sigma^2}\right) \left(\frac{\mu + \lambda}{\gamma\sigma^2}\right) = 0, \quad y \in [0, \log u(\lambda)/l(\lambda)], \quad (3.21)$$

$$w(0) = \frac{\mu - \lambda}{\gamma\sigma^2}, \quad (3.22)$$

$$w(\log(u(\lambda)/l(\lambda))) = \frac{\mu + \lambda}{\gamma\sigma^2}. \quad (3.23)$$

In other words, the correct value of λ is identified as the one for which the above first-order Riccati equation satisfies both the initial and the terminal value conditions. For a fixed ε , such a value is the solution of a scalar equation obtained from the explicit solution of the Riccati equation (cf. Lemma 3.1). For $\varepsilon \sim 0$, the asymptotic expansion of $\lambda(\varepsilon)$ follows from the implicit function theorem – and some patient calculations (see Lemma 3.2 below).

Now, one could argue that the advantage of the nonlinear, first-order equation (3.21) over the linear, second-order equation (3.18) is only marginal. In fact, the variable w has the additional advantage, albeit still hidden at this point, that it coincides with the optimal *shadow* risky portfolio weight (cf. Lemma 4.4), a fact that is hinted at by its boundary conditions. Furthermore, and as a result, for $\gamma = 1$ the equation (3.21) recovers the case of logarithmic utility (Gerhold et al., 2011b), while (3.18) does not.

3.3 Explicit formulas

Let us now show that the reduced value function w and the quantity λ are indeed well-defined. To this end, first determine, for a given small $\lambda > 0$, an explicit expression for the solution w of the ODE (3.21), complemented by the initial condition (3.22).

Lemma 3.1. *Let $0 < \mu/\gamma\sigma^2 \neq 1$. Then for sufficiently small $\lambda > 0$, the function*

$$w(\lambda, y) = \begin{cases} \frac{a(\lambda) \tanh[\tanh^{-1}(b(\lambda)/a(\lambda)) - a(\lambda)y] + (\frac{\mu}{\sigma^2} - \frac{1}{2})}{\gamma - 1}, & \text{if } \gamma \in (0, 1) \text{ and } \frac{\mu}{\gamma\sigma^2} < 1 \text{ or } \gamma > 1 \text{ and } \frac{\mu}{\gamma\sigma^2} > 1, \\ \frac{a(\lambda) \tan[\tan^{-1}(b(\lambda)/a(\lambda)) + a(\lambda)y] + (\frac{\mu}{\sigma^2} - \frac{1}{2})}{\gamma - 1}, & \text{if } \gamma > 1 \text{ and } \frac{\mu}{\gamma\sigma^2} \in \left(\frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{1}{\gamma}}, \frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{1}{\gamma}}\right), \\ \frac{a(\lambda) \coth[\coth^{-1}(b(\lambda)/a(\lambda)) - a(\lambda)y] + (\frac{\mu}{\sigma^2} - \frac{1}{2})}{\gamma - 1}, & \text{otherwise,} \end{cases}$$

with

$$a(\lambda) = \sqrt{\left|(\gamma - 1)\frac{\mu^2 - \lambda^2}{\gamma\sigma^4} - \left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2\right|} \quad \text{and} \quad b(\lambda) = \frac{1}{2} - \frac{\mu}{\sigma^2} + (\gamma - 1)\frac{\mu - \lambda}{\gamma\sigma^2},$$

is a local solution of

$$w'(y) + (1-\gamma)w^2(y) + \left(\frac{2\mu}{\sigma^2} - 1\right) w(y) - \frac{\mu^2 - \lambda^2}{\gamma\sigma^4} = 0, \quad w(0) = \frac{\mu - \lambda}{\gamma\sigma^2}. \quad (3.24)$$

Moreover, $y \mapsto w(\lambda, y)$ is increasing (resp. decreasing) for $\mu/\gamma\sigma^2 \in (0, 1)$ (resp. $\mu/\gamma\sigma^2 > 1$).

Proof. The first part of the assertion is easily verified by taking derivatives. The second follows by inspection of the explicit formulas. \square

Next, establish that the crucial constant λ , which determines both the no-trade region and the equivalent safe rate, is well-defined. For small transaction costs $\varepsilon \sim 0$, its asymptotics are readily computed by means of the implicit function theorem.

Lemma 3.2. *Let $0 < \mu/\gamma\sigma^2 \neq 1$ and $w(\lambda, \cdot)$ be defined as in Lemma 3.1, and set*

$$l(\lambda) = \frac{\mu - \lambda}{\gamma\sigma^2 - (\mu - \lambda)}, \quad u(\lambda) = \frac{1}{(1 - \varepsilon)} \frac{\mu + \lambda}{\gamma\sigma^2 - (\mu + \lambda)}.$$

Then, for sufficiently small $\varepsilon > 0$, there exists a unique solution λ of

$$w\left(\lambda, \log\left(\frac{u(\lambda)}{l(\lambda)}\right)\right) - \frac{\mu + \lambda}{\gamma\sigma^2} = 0. \quad (3.25)$$

As $\varepsilon \downarrow 0$, it has the asymptotics

$$\lambda = \gamma\sigma^2 \left(\frac{3}{4\gamma} \left(\frac{\mu}{\gamma\sigma^2} \right)^2 \left(1 - \frac{\mu}{\gamma\sigma^2} \right)^2 \right)^{1/3} \varepsilon^{1/3} + O(\varepsilon).$$

Proof. Write the boundary condition (3.25) as $f(\lambda, \varepsilon) = 0$, where:

$$f(\lambda, \varepsilon) = w(\lambda, \log(u(\lambda)/l(\lambda))) - \frac{\mu + \lambda}{\gamma\sigma^2}.$$

Of course, $f(0, 0) = 0$ corresponds to the frictionless case. The implicit function theorem then suggests that for sufficiently small ε there exists a unique zero $\lambda(\varepsilon)$ with the asymptotics $\lambda(\varepsilon) \sim -\varepsilon f_\varepsilon/f_\lambda$, but the difficulty is that $f_\lambda = 0$, because λ is not of order ε . Heuristic arguments (Shreve and Soner, 1994; Rogers, 2004) suggest that λ is of order $\varepsilon^{1/3}$. Thus, setting $\lambda = \delta^{1/3}$ and $\hat{f}(\delta, \varepsilon) = f(\delta^{1/3}, \varepsilon)$, and computing the derivatives of the explicit formula for $w(\lambda, x)$ (cf. Lemma 3.1) shows that:

$$\hat{f}_\varepsilon(0, 0) = -\frac{\mu(\mu - \gamma\sigma^2)}{\gamma^2\sigma^4}, \quad \hat{f}_\delta(0, 0) = \frac{4}{3\mu^2\sigma^2 - 3\gamma\mu\sigma^4}.$$

As a result:

$$\delta(\varepsilon) \sim -\frac{\hat{f}_\varepsilon(0, 0)}{\hat{f}_\delta(0, 0)}\varepsilon = \frac{3\mu^2(\mu - \gamma\sigma^2)^2}{4\gamma^2\sigma^2}\varepsilon \quad \text{whence} \quad \lambda(\varepsilon) \sim \left(\frac{3\mu^2(\mu - \gamma\sigma^2)^2}{4\gamma^2\sigma^2} \right)^{1/3} \varepsilon^{1/3}.$$

\square

Henceforth, consider small transaction costs $\varepsilon > 0$, and let λ denote the constant in Lemma 3.2. Moreover, set $w(y) = w(\lambda, y)$, $a = a(\lambda)$, $b = b(\lambda)$, and $u = u(\lambda)$, $l = l(\lambda)$. In all cases, the function w can be extended smoothly to an open neighborhood of $[0, \log(u/l)]$ (resp. $[\log(u/l), 0]$ if $\mu/\gamma\sigma^2 > 1$). By continuity, the ODE (3.24) then also holds at 0 and $\log(u/l)$; inserting the boundary conditions for w yields the following counterparts for the derivative w' :

Lemma 3.3. *Let $0 < \mu/\gamma\sigma^2 \neq 1$. Then, in all three cases,*

$$w'(0) = \frac{\mu - \lambda}{\gamma\sigma^2} - \left(\frac{\mu - \lambda}{\gamma\sigma^2} \right)^2, \quad w'\left(\log\left(\frac{u}{l}\right)\right) = \frac{\mu + \lambda}{\gamma\sigma^2} - \left(\frac{\mu + \lambda}{\gamma\sigma^2} \right)^2.$$

3.4 Discussion

The above heuristics offer a practical approach to portfolio choice problems with transaction costs, and can be adapted to accommodate additional model features. More importantly, they yield results that are robust to the model specification. In view of (3.20) and the asymptotics for λ in Lemma 3.2, the no-trade boundaries have the expansion:

$$\pi_{\pm} = \frac{\mu}{\gamma\sigma^2} \pm \left(\frac{3}{4\gamma} \left(\frac{\mu}{\gamma\sigma^2} \right)^2 \left(1 - \frac{\mu}{\gamma\sigma^2} \right)^2 \right)^{1/3} \varepsilon^{1/3} + O(\varepsilon). \quad (3.26)$$

This expansion coincides with the one obtained by Janeček and Shreve (2004) in the model with consumption. In other words, the long-run and the consumption models yield exactly the same solution at the leading order for small transaction costs. The expansions do differ at the second order, but such differences tend to have a modest effect for typical parameter values.

A major advantage of the long-run objective is the possibility to reduce the solution to a single algebraic equation for the parameter λ , in terms of which the free boundaries are found explicitly. In principle, one could attempt the same reduction in the consumption problem, by substituting equations (3.13) and (3.15) into (3.12). The result is a scalar equation for l in terms of $v(l)$, the value of the reduced value function at the trading boundary. Alas, the equation does not have an explicit solution. Also, the value of $v(l)$ is identified as the only one for which the solution to the differential equation (which also has no explicit solution) matches the analogous boundary condition at u . The situation is disappointingly more complicated than (3.19), which immediately identifies both boundaries in terms of a single parameter. In summary, the consumption problem yields a solution which is strikingly similar to the long-run problem, but in a much less tractable setting. Vice versa, the long-run solution provides a tractable first-order approximation to the consumption problem.

In the same vein, the long-run optimal portfolio is not far from optimal for utility maximization with terminal wealth. Indeed, Gerhold et al. (2011a) show that the wealth corresponding to the long-run optimal portfolio matches the value function for any *finite* horizon T (Bichuch, 2012) at the leading order $\varepsilon^{2/3}$ for small transaction costs ε (compare Theorem 4.8 below). Hence, finite horizons – like consumption – only have a second-order effect on portfolio choice with transaction costs.

To apply the heuristic steps above to more complex problems with transaction costs, it is worth distinguishing the aspects that are special to the specific problem at hand from the ones that are flexible enough to be useful in other models. First, in general one cannot expect that a single, simple equation like (3.19) identifies both free boundaries. But the same argument that leads to this equation (the substitution of the boundary and smooth pasting conditions into the HJB equation) will generally lead in a long-run problem to some scalar equation for each boundary, in terms of the equivalent safe rate β of the problem. Such equations may be solved explicitly (as in the case of (3.19)) or not, but in the latter case an asymptotic solution will still be available, expanding the scalar equation around the frictionless values of (π, β) .

Second, the reduced HJB equation may not be autonomous or have an explicit solution, which are two special features of (3.21). If the equation is not autonomous, the Cauchy problem cannot be started at some arbitrary point (zero in the previous example) without a further change of variable. If the free boundaries admit explicit solutions in terms of β , sometimes a careful choice of notation can lead to a simple expression for at least one boundary, which is a natural choice for the starting point of the Cauchy problem. The correct value of β is then identified as the one for which the remaining boundary condition is satisfied. Even if the differential equation has an explicit solution,

this condition in general involves a scalar equation that cannot be solved explicitly. Regardless of an explicit formula, asymptotic expansions can be derived by substituting a series expansion for w in the differential equation.

4 Shadow Prices and Verification

We justify the heuristic arguments in the previous section by reducing the portfolio choice problem with transaction costs to another portfolio choice problem, without transaction costs. To do so, the bid and ask prices are replaced by a single “shadow price” \tilde{S}_t evolving within the bid-ask spread, which yields the same optimal policy and utility. Evidently, *any* frictionless market extension with values in the bid-ask spread leads to more favorable terms of trade than the original market with transaction costs. To achieve equality, the particularly unfavorable shadow price must match the trading prices whenever its optimal policy transacts. The latter is then also feasible and in turn optimal in the original market with transaction costs, motivating the following notion.

Definition 4.1. *A shadow price is a frictionless price process \tilde{S}_t evolving within the bid-ask spread $((1 - \varepsilon)S_t \leq \tilde{S}_t \leq S_t$ a.s. for all t), such that there is an optimal strategy for \tilde{S}_t which is of finite variation and entails buying only when the shadow price \tilde{S}_t equals the ask price S_t , and selling only when \tilde{S}_t equals the bid price $(1 - \varepsilon)S_t$.*

Once a candidate for such a shadow price is identified, long-run verification results for frictionless models (cf. Guasoni and Robertson (2012)) deliver the optimality of the guessed policy.

4.1 Derivation of a Candidate Shadow Price

With a smooth candidate value function at hand, a candidate shadow price is identified as follows. By definition, trading the shadow price should not allow the investor to outperform the original market with transaction costs. In particular, if \tilde{S}_t is the value of the shadow price at time t , then allowing the frictional investor to carry out a single trade at time t at this *frictionless* price should not allow her to increase her utility. A trade of ν risky shares at the frictionless price \tilde{S}_t moves the investor’s safe position X_t to $X_t - \nu\tilde{S}_t$ and her risky position (valued at the ask price S_t) from Y_t to $Y_t + \nu S_t$. Then – recalling that the second and third arguments of the candidate value functions V from the previous section were precisely the investor’s safe and risky positions – the requirement that such a trade does not increase the investor’s utility is tantamount to:

$$V(t, X_t - \nu\tilde{S}_t, Y_t + \nu S_t) \leq V(t, X_t, Y_t), \quad \forall \nu \in \mathbb{R}.$$

A Taylor expansion of the left-hand side for small ν then implies that $-\nu\tilde{S}_t V_x + \nu S_t V_y \leq 0$. Since this inequality has to hold both for positive and negative values of ν , it implies that

$$\tilde{S}_t = \frac{V_y}{V_x} S_t. \tag{4.1}$$

That is, the multiplicative deviation of the shadow price from the ask price should be the marginal rate of substitution of risky for safe assets for the optimal frictional investor. In particular, this formula immediately yields a candidate shadow price, once a smooth candidate value function has been identified. For the long-run problem, we derived the following candidate value function in the previous section:

$$V(t, X_t, Y_t) = e^{-(1-\gamma)(r+\beta)t} (X_t)^{1-\gamma} e^{(1-\gamma) \int_0^{\log(Y_t/X_t)} w(y) dy}.$$

Using this equality to calculate the partial derivatives in (4.1), the candidate shadow price becomes:

$$\tilde{S}_t = \frac{w(\Upsilon_t)}{le^{\Upsilon_t}(1-w(\Upsilon_t))} S_t, \quad (4.2)$$

where $\Upsilon_t = \log(X_t/lX_t^0)$ denotes the logarithm of the stock-cash ratio, centered in its value at the lower buying boundary l . If this candidate is indeed the right one, then its optimal strategy and value function should coincide with their frictional counterparts derived heuristically above. In particular, the optimal risky fraction $\tilde{\pi}_t$ should correspond to the same numbers φ_t^0 and φ_t of safe and risky shares, but now measured in terms of \tilde{S}_t instead of the ask price S_t . As a consequence:

$$\tilde{\pi}_t = \frac{\varphi_t \tilde{S}_t}{\varphi_t^0 S_t^0 + \varphi_t \tilde{S}_t} = \frac{\varphi_t S_t \frac{w(\Upsilon_t)}{le^{\Upsilon_t}(1-w(\Upsilon_t))}}{\varphi_t^0 S_t^0 + \varphi_t S_t \frac{w(\Upsilon_t)}{le^{\Upsilon_t}(1-w(\Upsilon_t))}} = \frac{\frac{w(\Upsilon_t)}{1-w(\Upsilon_t)}}{1 + \frac{w(\Upsilon_t)}{1-w(\Upsilon_t)}} = w(\Upsilon_t), \quad (4.3)$$

where, for the third equality, we have used that the optimal frictional stock-cash ratio $\varphi_t S_t / \varphi_t^0 S_t^0$ equals le^{Υ_t} by definition of Υ_t . We now turn to the corresponding value function \tilde{V} . By the definition of shadow price, it should coincide with its frictional counterpart V . In the frictionless case, it is more convenient to factor out the total wealth $\tilde{X}_t = \varphi_t^0 S_t^0 + \varphi_t \tilde{S}_t$ (in terms of the frictionless risky price \tilde{S}_t) instead of the safe position $X_t = \varphi_t^0 S_t^0$, giving

$$\tilde{V}(t, \tilde{X}_t, \Upsilon_t) = V(t, X_t, Y_t) = e^{-(1-\gamma)(r+\beta)t} \tilde{X}_t^{1-\gamma} \left(\frac{X_t}{\tilde{X}_t} \right)^{1-\gamma} e^{(1-\gamma) \int_0^{\Upsilon_t} w(y) dy}.$$

Since $X_t/\tilde{X}_t = 1 - w(\Upsilon_t)$ by definition of \tilde{S}_t , one can rewrite the last two factors as

$$\begin{aligned} \left(\frac{X_t}{\tilde{X}_t} \right)^{1-\gamma} e^{(1-\gamma) \int_0^{\Upsilon_t} w(y) dy} &= \exp \left((1-\gamma) \left[\log(1-w(\Upsilon_t)) + \int_0^{\Upsilon_t} w(y) dy \right] \right) \\ &= (1-w(0))^{\gamma-1} \exp \left((1-\gamma) \int_0^{\Upsilon_t} \left(w(y) - \frac{w'(y)}{1-w(y)} \right) dy \right). \end{aligned}$$

Then, setting $\tilde{w} = w - \frac{w'}{1-w}$, the candidate long-run value function for \tilde{S} becomes

$$\tilde{V}(t, \tilde{X}_t, \Upsilon_t) = e^{-(1-\gamma)(r+\beta)t} \tilde{X}_t^{1-\gamma} e^{(1-\gamma) \int_0^{\Upsilon_t} \tilde{w}(y) dy} (1-w(0))^{\gamma-1}.$$

Starting from the candidate value function and optimal policy for \tilde{S} , we can now proceed to verify that they are indeed optimal for \tilde{S}_t , by adapting the argument from Guasoni and Robertson (2012). But before we do that, we have to construct the respective shadow processes.

4.2 Construction of the Shadow Price

The above heuristic arguments suggest that the optimal stock-cash ratio $Y_t/X_t = \varphi_t S_t / \varphi_t^0 S_t^0$ should take values in the interval $[l, u]$. Hence, $\Upsilon_t = \log(Y_t/lX_t)$ should be $[0, \log(u/l)]$ -valued if the lower trading boundary l for the stock-cash ratio X_t/X_t^0 is positive. If the investor shorts the safe asset to leverage her risky position, the stock-cash ratio becomes negative. In the frictionless case, and also for small transaction costs, this happens if the Merton proportion $\mu/\gamma\sigma^2$ is bigger than 1. Then, the trading boundaries $l \leq u$ are both negative, so that the centered log-stock-cash ratio Υ_t should take values in $[\log(u/l), 0]$. In both cases, trading should only take place when the stock-cash ratio reaches the boundaries of this region. Hence, the numbers of safe and risky units φ_t^0 and φ_t should remain constant and $\Upsilon_t = \log(\varphi_t/l\varphi_t^0) + \log(S_t/S_t^0)$ should follow a Brownian motion with drift as

long as Υ_t moves in $(0, \log(u/l))$ (resp. in $(\log(u/l), 0)$ if $\mu/\gamma\sigma^2 > 1$). This motivates to *define* the process Υ_t as reflected Brownian motion:

$$d\Upsilon_t = (\mu - \sigma^2/2)dt + \sigma dW_t + dL_t - dU_t, \quad \Upsilon_0 \in [0, \log(u/l)], \quad (4.4)$$

for continuous, adapted local time processes L and U which are nondecreasing (resp. nonincreasing if $\mu/\gamma\sigma^2 > 1$) and increase (resp. decrease if $\mu/\gamma\sigma^2 > 1$) only on the sets $\{\Upsilon_t = 0\}$ and $\{\Upsilon_t = \log(u/l)\}$, respectively. Starting from this process, whose existence is a classical result of Skorohod (1962), the process \tilde{S} is defined in accordance with (4.2):

Lemma 4.2. *Let $(\xi^0, \xi) \in \mathbb{R}_+^2$ be the investor's initial endowment in units of the safe and risky asset. Define*

$$y = \begin{cases} 0, & \text{if } l\xi^0 S_0^0 \geq \xi S_0, \\ \log(u/l), & \text{if } u\xi^0 S_0^0 \leq \xi S_0, \\ \log[\xi S_0/(\xi^0 S_0^0 l)], & \text{otherwise,} \end{cases} \quad (4.5)$$

and let Υ be defined as in (4.4), starting at $\Upsilon_0 = y$. Then, $\tilde{S} = S \frac{w(\Upsilon)}{le^{\Upsilon}(1-w(\Upsilon))}$, with w as in Lemma 3.1, has the dynamics

$$d\tilde{S}(\Upsilon_t)/\tilde{S}(\Upsilon_t) = (\tilde{\mu}(\Upsilon_t) + r)dt + \tilde{\sigma}(\Upsilon_t)dW_t,$$

where $\tilde{\mu}(\cdot)$ and $\tilde{\sigma}(\cdot)$ are defined as

$$\tilde{\mu}(y) = \frac{\sigma^2 w'(y)}{w(y)(1-w(y))} \left(\frac{w'(y)}{1-w(y)} - (1-\gamma)w(y) \right), \quad \tilde{\sigma}(y) = \frac{\sigma w'(y)}{w(y)(1-w(y))}.$$

Moreover, the process \tilde{S} takes values within the bid-ask spread $[(1-\varepsilon)S, S]$.

Note that the first two cases in (4.5) arise if the initial stock-cash ratio $\xi S_0/(\xi^0 S_0^0)$ lies outside of the interval $[l, u]$. Then, a jump from the initial position $(\varphi_{0-}^0, \varphi_{0-}) = (\xi^0, \xi)$ to the nearest boundary value of $[l, u]$ is required. This transfer requires the purchase resp. sale of the risky asset and hence the initial price \tilde{S}_0 is defined to match the buying resp. selling price of the risky asset.

Proof. The dynamics of \tilde{S}_t result from Itô's formula, the dynamics of Υ_t , and the identity

$$w''(y) = 2(\gamma - 1)w'(y)w(y) - (2\mu/\sigma^2 - 1)w'(y), \quad (4.6)$$

obtained by differentiating the ODE (3.24) for w with respect to x . Therefore it remains to show that \tilde{S}_t indeed takes values in the bid-ask spread $[(1-\varepsilon)S_t, S_t]$. To this end, notice that – in view of the ODE (3.24) for w – the derivative of the function $g(y) := w(y)/le^y(1-w(y))$ is given by

$$g'(y) = \frac{w'(y) - w(y) + w^2(y)}{le^y(1-w(y))^2} = \frac{\gamma(w^2 - 2\frac{\mu}{\gamma\sigma^2}w) + (\mu^2 - \lambda^2)/\gamma\sigma^4}{le^y(1-w(y))^2}.$$

Due to the boundary conditions for w , the derivative g' vanishes at 0 and $\log(u/l)$. Differentiating its numerator gives $2\gamma w'(y)(w(y) - \frac{\mu}{\gamma\sigma^2})$. For $\frac{\mu}{\gamma\sigma^2} \in (0, 1)$ (resp. $\frac{\mu}{\gamma\sigma^2} > 1$), w is increasing from $\frac{\mu-\lambda}{\gamma\sigma^2} < \frac{\mu}{\gamma\sigma^2}$ to $\frac{\mu+\lambda}{\gamma\sigma^2} > \frac{\mu}{\gamma\sigma^2}$ on $[0, \log(u/l)]$ (resp. decreasing from $\frac{\mu+\lambda}{\gamma\sigma^2}$ to $\frac{\mu-\lambda}{\gamma\sigma^2}$ on $[\log(u/l), 0]$); hence, w' is nonnegative (resp. nonpositive). Moreover, g' starts at zero for $y = 0$ (resp. $\log(u/l)$), then decreases (resp. increases), and eventually starts increasing (resp. decreasing) again, until it reaches level zero again for $y = \log(u/l)$ (resp. $y = 0$). In particular, g' is nonpositive (resp. nonnegative), so that g is decreasing on $[0, \log(u/l)]$ (resp. increasing on $[\log(u/l), 0]$ for $\frac{\mu}{\gamma\sigma^2} > 1$). Taking into account that $g(0) = 1$ and $g(\log(u/l)) = 1 - \varepsilon$, by the boundary conditions for w and the definition of u and l in Lemma 3.2, the proof is now complete. \square

4.3 Verification

The long-run optimality of the candidate risky weight $\tilde{\pi}(\Upsilon_t) = w(\Upsilon_t)$ from (4.3) in the frictionless market with price process \tilde{S}_t can now be verified by adapting the argument in Guasoni and Robertson (2012). The first step is to determine finite-horizon bounds, which provide lower and upper estimates for the maximal expected utility on any horizon T , by focusing on the values of the candidate long-run optimal policy and long-run optimal martingale measure.

These bounds are based on the concept of the (long-run) myopic probability, the hypothetical probability measure under which a logarithmic investor would adopt the same policy as the original power investor under the physical probability. The advantage of this probability is to decompose expected *power* utility (and its dual) into a long-run component *times* a transient component. This decomposition is similar in spirit to the separation of *logarithmic* utility into a long-run component *plus* a transitory component. To see the analogy, consider the logarithmic utility of a portfolio $\pi(\Theta_t)$ traded in a frictionless market with expected excess return $\tilde{\mu}(\Theta_t)$ and volatility $\tilde{\sigma}(\Theta_t)$ driven by some state variable Θ_t :

$$\log \tilde{X}_T^\pi = x + \int_0^T \left(\tilde{\mu}(\Theta_t)\pi(\Theta_t) - \frac{\tilde{\sigma}^2(\Theta_t)}{2}\pi^2(\Theta_t) \right) dt + \int_0^T \tilde{\sigma}(\Theta_t)\pi(\Theta_t)dW_t.$$

Now, if Θ_t follows an autonomous diffusion $d\Theta_t = b(\Theta_t)dt + dW_t$, the above stochastic integral can be replaced by applying Itô's formula to the function $\Pi(y) = \int_0^y \tilde{\sigma}(x)\pi(x)dx$:

$$\Pi(\Theta_T) - \Pi(\Theta_0) = \int_0^T \left(\tilde{\sigma}(\Theta_t)\pi(\Theta_t)b(\Theta_t) + \frac{1}{2}(\tilde{\sigma}\pi)'(\Theta_t) \right) dt + \int_0^T \tilde{\sigma}(\Theta_t)\pi(\Theta_t)dW_t.$$

Indeed, solving the second equation for the stochastic integral, and plugging it into the first equation yields:

$$\log \tilde{X}_T^\pi = x + \int_0^T \left((\tilde{\mu}(\Theta_t) - \tilde{\sigma}(\Theta_t)b(\Theta_t))\pi(\Theta_t) - \frac{\tilde{\sigma}^2(\Theta_t)}{2}\pi^2(\Theta_t) - \frac{(\tilde{\sigma}\pi)'(\Theta_t)}{2} \right) dt + (\Pi(\Theta_T) - \Pi(\Theta_0)).$$

This decomposes the logarithmic utility into an integral, which represents the long-run component, and a residual transitory term, which depends only on the initial and terminal values of the state variable. If the function Π is integrable with respect to the invariant measure of Θ , the contribution of the transitory component to the equivalent safe rate $\frac{1}{T}E[\log \tilde{X}_T^\pi]$ is negligible for long horizons.

The myopic probability is key to perform a similar decomposition with power utility. Again, denote by $\tilde{\mu}$ the risky asset's drift under the original measure, and by $\hat{\mu}$ its counterpart under the myopic probability; the corresponding volatility $\tilde{\sigma}$ of course has to be the same under both equivalent measures. With logarithmic utility, the optimal portfolio is $\hat{\pi}_t = \hat{\mu}_t/\tilde{\sigma}_t^2$ even if $\hat{\mu}_t$ and $\tilde{\sigma}_t$ are stochastic (Merton, 1971). As the definition of the myopic probability requires that the corresponding log-optimal portfolio $\hat{\pi}_t$ coincides with the optimal portfolio $\tilde{\pi}_t$ for power utility under the original probability, Girsanov's theorem dictates that the measure change from the original to the myopic probability is governed by the stochastic exponential of $\int_0^T (-\frac{\tilde{\mu}}{\tilde{\sigma}} + \tilde{\sigma}\tilde{\pi})dW_t$. This measure change shifts the asset's drift by the same amount, times $\tilde{\sigma}$, thereby yielding a myopic drift of $\tilde{\sigma}^2\tilde{\pi}$, which yields the same optimal policy. Given this guess for the myopic probability, the finite-horizon bounds follow by routine calculations carried out in the proof of the following lemma:

Lemma 4.3. *For a fixed time horizon $T > 0$, let $\beta = \frac{\mu^2 - \lambda^2}{2\gamma\sigma^2}$ and let the function w be defined as in Lemma 3.1. Then, for the the shadow payoff \tilde{X}_T corresponding to the policy $\tilde{\pi}(\Upsilon_t) = w(\Upsilon_t)$ and*

the shadow discount factor $\tilde{M}_T = e^{-rT} \mathcal{E}(-\int_0^T \frac{\tilde{\mu}}{\tilde{\sigma}} dW_t)_T$, the following bounds hold true:

$$E[\tilde{X}_T^{1-\gamma}] = \tilde{X}_0^{1-\gamma} e^{(1-\gamma)(r+\beta)T} \hat{E}[e^{(1-\gamma)(\tilde{q}(\Upsilon_0)-\tilde{q}(\Upsilon_T))}], \quad (4.7)$$

$$E\left[\tilde{M}_T^{1-\frac{1}{\gamma}}\right]^\gamma = e^{(1-\gamma)(r+\beta)T} \hat{E}\left[e^{(\frac{1}{\gamma}-1)(\tilde{q}(\Upsilon_0)-\tilde{q}(\Upsilon_T))}\right]^\gamma, \quad (4.8)$$

where $\tilde{q}(y) := \int_0^y (w(z) - \frac{w'(z)}{1-w(z)}) dz$ and $\hat{E}[\cdot]$ denotes the expectation with respect to the myopic probability \hat{P} , defined by

$$\frac{d\hat{P}}{dP} = \exp\left(\int_0^T \left(-\frac{\tilde{\mu}(\Upsilon_t)}{\tilde{\sigma}(\Upsilon_t)} + \tilde{\sigma}(\Upsilon_t)\tilde{\pi}(\Upsilon_t)\right) dW_t - \frac{1}{2} \int_0^T \left(-\frac{\tilde{\mu}(\Upsilon_t)}{\tilde{\sigma}(\Upsilon_t)} + \tilde{\sigma}(\Upsilon_t)\tilde{\pi}(\Upsilon_t)\right)^2 dt\right).$$

Proof. First note that $\tilde{\mu}, \tilde{\sigma}$, and w are functions of Υ_t , but the argument is omitted throughout to ease notation. Now, to prove (4.7), notice that the frictionless shadow wealth process \tilde{X}_t with dynamics $\frac{d\tilde{X}_t}{\tilde{X}_t} = w \frac{d\tilde{S}_t}{\tilde{S}_t} + (1-w) \frac{dS_t^0}{S_t^0}$ satisfies:

$$\tilde{X}_T^{1-\gamma} = \tilde{X}_0^{1-\gamma} e^{(1-\gamma) \int_0^T (r + \tilde{\mu}w - \frac{\tilde{\sigma}^2}{2} w^2) dt + (1-\gamma) \int_0^T \tilde{\sigma} w dW_t}.$$

Hence:

$$\tilde{X}_T^{1-\gamma} = \tilde{X}_0^{1-\gamma} \frac{d\hat{P}}{dP} e^{\int_0^T ((1-\gamma)(r + \tilde{\mu}w - \frac{\tilde{\sigma}^2}{2} w^2) + \frac{1}{2}(-\frac{\tilde{\mu}}{\tilde{\sigma}} + \tilde{\sigma}w)^2) dt + \int_0^T ((1-\gamma)\tilde{\sigma}w - (-\frac{\tilde{\mu}}{\tilde{\sigma}} + \tilde{\sigma}w)) dW_t}.$$

Inserting the definitions of $\tilde{\mu}$ and $\tilde{\sigma}$, the second integrand simplifies to $(1-\gamma)\sigma(\frac{w'}{1-w} - w)$. Similarly, the first integrand reduces to $(1-\gamma)(r + \frac{\sigma^2}{2}(\frac{w'}{1-w})^2 - (1-\gamma)\sigma^2 \frac{w'w}{1-w} + (1-\gamma)\frac{\sigma^2}{2}w^2)$. In summary:

$$\tilde{X}_T^{1-\gamma} = \tilde{X}_0^{1-\gamma} \frac{d\hat{P}}{dP} e^{(1-\gamma) \int_0^T (r + \frac{\sigma^2}{2}(\frac{w'}{1-w})^2 - (1-\gamma)\sigma^2 \frac{w'w}{1-w} + (1-\gamma)\frac{\sigma^2}{2}w^2) dt + (1-\gamma) \int_0^T \sigma(\frac{w'}{1-w} - w) dW_t}. \quad (4.9)$$

The boundary conditions for w and w' imply $w(0) - \frac{w'(0)}{1-w(0)} = w(\log(u/l)) - \frac{w'(\log(u/l))}{1-w(\log(u/l))} = 0$; hence, Itô's formula yields that the local time terms vanish in the dynamics of $\tilde{q}(\Upsilon_t)$:

$$\tilde{q}(\Upsilon_T) - \tilde{q}(\Upsilon_0) = \int_0^T \left(\mu - \frac{\sigma^2}{2}\right) \left(w - \frac{w'}{1-w}\right) + \frac{\sigma^2}{2} \left(w' - \frac{w''(1-w) + w'^2}{(1-w)^2}\right) dt + \int_0^T \sigma \left(w - \frac{w'}{1-w}\right) dW_t. \quad (4.10)$$

Substituting the second derivative w'' according to the ODE (4.6) and using the resulting identity to replace the stochastic integral in (4.9) yields

$$\tilde{X}_T^{1-\gamma} = \tilde{X}_0^{1-\gamma} \frac{d\hat{P}}{dP} e^{(1-\gamma) \int_0^T (r + \frac{\sigma^2}{2} w' + (1-\gamma) \frac{\sigma^2}{2} w^2 + (\mu - \frac{\sigma^2}{2}) w) dt} e^{(1-\gamma)(\tilde{q}(\Upsilon_0) - \tilde{q}(\Upsilon_T))}.$$

After inserting the ODE (3.24) for w , the first bound thus follows by taking the expectation.

The argument for the second bound is similar. Plugging in the definitions of $\tilde{\mu}$ and $\tilde{\sigma}$, the shadow discount factor $\tilde{M}_T = e^{-rT} \mathcal{E}(-\int_0^T \frac{\tilde{\mu}}{\tilde{\sigma}} dW_t)_T$ and the myopic probability \hat{P} satisfy:

$$\begin{aligned} \tilde{M}_T^{1-\frac{1}{\gamma}} &= e^{\frac{1-\gamma}{\gamma} \int_0^T \frac{\tilde{\mu}}{\tilde{\sigma}} dW_t + \frac{1-\gamma}{\gamma} \int_0^T (r + \frac{\tilde{\mu}^2}{2\tilde{\sigma}^2}) dt} \\ &= \frac{d\hat{P}}{dP} e^{\frac{1-\gamma}{\gamma} \int_0^T (\frac{\tilde{\mu}}{\tilde{\sigma}} - \frac{\gamma}{1-\gamma}(-\frac{\tilde{\mu}}{\tilde{\sigma}} + \tilde{\sigma}w)) dW_t + \frac{1-\gamma}{\gamma} \int_0^T (r + \frac{\tilde{\mu}^2}{2\tilde{\sigma}^2} + \frac{\gamma}{2(1-\gamma)}(-\frac{\tilde{\mu}}{\tilde{\sigma}} + \tilde{\sigma}w)^2) dt} \\ &= \frac{d\hat{P}}{dP} e^{\frac{1-\gamma}{\gamma} \int_0^T \sigma(\frac{w'}{1-w} - w) dW_t + \frac{1-\gamma}{\gamma} \int_0^T (r + \frac{\sigma^2}{2}(\frac{w'}{1-w})^2 - (1-\gamma)\sigma^2 \frac{w'w}{1-w} + (1-\gamma)\frac{\sigma^2}{2}w^2) dt}. \end{aligned}$$

Again replace the stochastic integral using (4.10) and the ODE (4.6), obtaining

$$\tilde{M}_T^{1-\frac{1}{\gamma}} = \frac{d\hat{P}}{dP} e^{\frac{1-\gamma}{\gamma} \int_0^T (r + \frac{\sigma^2}{2} w' + (1-\gamma) \frac{\sigma^2}{2} w^2 + (\mu - \frac{\sigma^2}{2}) w) dt} e^{\frac{1-\gamma}{\gamma} (\tilde{q}(\Upsilon_0) - \tilde{q}(\Upsilon_T))}.$$

Inserting the ODE (3.24) for w , taking the expectation, and raising it to power γ , the second bound follows. \square

With the finite horizon bounds at hand, it is now straightforward to establish that the policy $\tilde{\pi}(\Upsilon_t)$ is indeed long-run optimal in the frictionless market with price \tilde{S}_t .

Lemma 4.4. *Let $0 < \mu/\gamma\sigma^2 \neq 1$ and let w be defined as in Lemma 3.1. Then, the risky weight $\tilde{\pi}(\Upsilon_t) = w(\Upsilon_t)$ is long-run optimal with equivalent safe rate $r + \beta$ in the frictionless market with price process \tilde{S}_t . The corresponding wealth process (in terms of \tilde{S}_t), and the numbers of safe and risky units are given by*

$$\begin{aligned} \tilde{X}_t &= (\xi^0 S_0^0 + \xi \tilde{S}_0) \mathcal{E} \left(\int_0^t (r + w(\Upsilon_s) \tilde{\mu}(\Upsilon_s)) ds + \int_0^t w(\Upsilon_s) \tilde{\sigma}(\Upsilon_s) dW_s \right)_t, \\ \varphi_{0-} &= \xi, \quad \varphi_t = w(\Upsilon_t) \tilde{X}_t / \tilde{S}_t \quad \text{for } t \geq 0, \\ \varphi_{0-}^0 &= \xi^0, \quad \varphi_t^0 = (1 - w(\Upsilon_t)) \tilde{X}_t / S_t^0 \quad \text{for } t \geq 0. \end{aligned}$$

Proof. The formulas for the wealth process and the corresponding numbers of safe and risky units follow directly from the standard frictionless definitions. Now let \tilde{M}_t be the shadow discount factor from Lemma 4.3. Then, standard duality arguments for power utility (cf. Lemma 5 in Guasoni and Robertson (2012)) imply that the shadow payoff \tilde{X}_t^ϕ corresponding to *any* admissible strategy ϕ_t satisfies the inequality

$$E \left[(\tilde{X}_T^\phi)^{1-\gamma} \right]^{\frac{1}{1-\gamma}} \leq E \left[\tilde{M}_T^{\frac{\gamma-1}{\gamma}} \right]^{\frac{\gamma}{1-\gamma}}. \quad (4.11)$$

This inequality in turn yields the following upper bound, valid for any admissible strategy ϕ_t in the frictionless market with shadow price \tilde{S}_t :

$$\liminf_{T \rightarrow \infty} \frac{1}{(1-\gamma)T} \log E \left[(\tilde{X}_T^\phi)^{1-\gamma} \right] \leq \liminf_{T \rightarrow \infty} \frac{\gamma}{(1-\gamma)T} \log E \left[\tilde{M}_T^{\frac{\gamma-1}{\gamma}} \right]. \quad (4.12)$$

Since the function \tilde{q} is bounded on the compact support of Υ_t , the second bound in Lemma 4.3 implies that the right-hand side equals $r + \beta$. Likewise, the first bound in the same lemma implies that the shadow payoff \tilde{X}_t (corresponding to the policy φ_t) attains this upper bound, concluding the proof. \square

The next Lemma establishes that the candidate \tilde{S}_t is indeed a shadow price.

Lemma 4.5. *Let $0 < \mu/\gamma\sigma^2 \neq 1$. Then, the number of shares $\varphi_t = w(\Upsilon_t) \tilde{X}_t / \tilde{S}_t$ in the portfolio $\tilde{\pi}(\Upsilon_t)$ in Lemma 4.4 has the dynamics*

$$\frac{d\varphi_t}{\varphi_t} = \left(1 - \frac{\mu - \lambda}{\gamma\sigma^2} \right) dL_t - \left(1 - \frac{\mu + \lambda}{\gamma\sigma^2} \right) dU_t. \quad (4.13)$$

Thus, φ_t increases only when $\Upsilon_t = 0$, that is, when \tilde{S}_t equals the ask price, and decreases only when $\Upsilon_t = \log(u/l)$, that is, when \tilde{S}_t equals the bid price.

Proof. Itô's formula and the ODE (4.6) yield

$$dw(\Upsilon_t) = -(1 - \gamma)\sigma^2 w'(\Upsilon_t)w(\Upsilon_t)dt + \sigma w'(\Upsilon_t)dW_t + w'(\Upsilon_t)(dL_t - dU_t).$$

Integrating $\varphi_t = w(\Upsilon_t)\tilde{X}_t/\tilde{S}_t$ by parts twice, inserting the dynamics of $w(\Upsilon_t)$, \tilde{X}_t , \tilde{S}_t , and simplifying yields:

$$\frac{d\varphi_t}{\varphi_t} = \frac{w'(\Upsilon_t)}{w(\Upsilon_t)}d(L_t - U_t).$$

Since L_t and U_t only increase (resp. decrease when $\mu/\gamma\sigma^2 > 1$) on $\{\Upsilon_t = 0\}$ and $\{\Upsilon_t = \log(u/l)\}$, respectively, the assertion now follows from the boundary conditions for w and w' . \square

The optimal growth rate for any frictionless price within the bid-ask spread must be greater or equal than in the original market with bid-ask process $((1 - \varepsilon)S_t, S_t)$, because the investor trades at more favorable prices. For a *shadow price*, there is an optimal strategy that only entails buying (resp. selling) stocks when \tilde{S}_t coincides with the ask- resp. bid price. Hence, this strategy yields the same payoff when executed at bid-ask prices, and thus is also optimal in the original model with transaction costs. The corresponding equivalent safe rate must also be the same, since the difference due to the liquidation costs vanishes as the horizon grows in (3.3):

Proposition 4.6. *For a sufficiently small spread ε , the strategy (φ_t^0, φ_t) from Lemma 4.4 is also long-run optimal in the original market with transaction costs, with the same equivalent safe rate.*

Proof. As φ_t only increases (resp. decreases) when $\tilde{S}_t = S_t$ (resp. $\tilde{S}_t = (1 - \varepsilon)S_t$), the strategy (φ_t^0, φ_t) is also self-financing for the bid-ask process $((1 - \varepsilon)S_t, S_t)$. Since $S_t \geq \tilde{S}_t \geq (1 - \varepsilon)S_t$ and the number φ_t of risky shares is always positive, it follows that

$$\varphi_t^0 S_t^0 + \varphi_t \tilde{S}_t \geq \varphi_t^0 S_t^0 + \varphi_t^+(1 - \varepsilon)S_t - \varphi_t^- S_t \geq (1 - \frac{\varepsilon}{1 - \varepsilon})\tilde{\pi}(Y_t)(\varphi_t^0 S_t^0 + \varphi_t \tilde{S}_t). \quad (4.14)$$

The shadow risky fraction $\tilde{\pi}(\Upsilon_t) = w(\Upsilon_t)$ is bounded from above by $(\mu + \lambda)/\gamma\sigma^2 = \mu/\gamma\sigma^2 + O(\varepsilon^{1/3})$. For a sufficiently small spread ε , the strategy (φ_t^0, φ_t) is therefore also admissible for $((1 - \varepsilon)S_t, S_t)$. Moreover, (4.14) then also yields

$$\begin{aligned} \liminf_{T \rightarrow \infty} \frac{1}{(1 - \gamma)T} \log E [(\varphi_T^0 S_T^0 + \varphi_T^+(1 - \varepsilon)S_T - \varphi_T^- S_T)^{1 - \gamma}] \\ = \liminf_{T \rightarrow \infty} \frac{1}{(1 - \gamma)T} \log E [(\varphi_T^0 S_T^0 + \varphi_T \tilde{S}_T)^{1 - \gamma}], \end{aligned} \quad (4.15)$$

that is, (φ_t^0, φ_t) has the same growth rate, either with \tilde{S}_t or with $[(1 - \varepsilon)S_t, S_t]$.

For any admissible strategy (ψ_t^0, ψ_t) for the bid-ask spread $[(1 - \varepsilon)S_t, S_t]$, set $\tilde{\psi}_t^0 = \psi_{0-}^0 - \int_0^t \tilde{S}_s/S_s^0 d\psi_s$. Then, $(\tilde{\psi}_t^0, \psi_t)$ is a self-financing trading strategy for \tilde{S}_t with $\tilde{\psi}_t^0 \geq \psi_t^0$. Together with $\tilde{S}_t \in [(1 - \varepsilon)S_t, S_t]$, the long-run optimality of (φ_t^0, φ_t) for \tilde{S}_t , and (4.15), it follows that:

$$\begin{aligned} \liminf_{T \rightarrow \infty} \frac{1}{T} \frac{1}{(1 - \gamma)} \log E [(\psi_T^0 S_T^0 + \psi_T^+(1 - \varepsilon)S_T - \psi_T^- S_T)^{1 - \gamma}] \\ \leq \liminf_{T \rightarrow \infty} \frac{1}{T} \frac{1}{(1 - \gamma)} \log E [(\tilde{\psi}_T^0 S_T^0 + \psi_T \tilde{S}_T)^{1 - \gamma}] \\ \leq \liminf_{T \rightarrow \infty} \frac{1}{T} \frac{1}{(1 - \gamma)} \log E [(\varphi_T^0 S_T^0 + \varphi_T \tilde{S}_T)^{1 - \gamma}] \\ = \liminf_{T \rightarrow \infty} \frac{1}{T} \frac{1}{(1 - \gamma)} \log E [(\varphi_T^0 S_T^0 + \varphi_T^+(1 - \varepsilon)S_T - \varphi_T^- S_T)^{1 - \gamma}]. \end{aligned}$$

Hence (φ_t^0, φ_t) is also long-run optimal for $((1 - \varepsilon)S_t, S_t)$. \square

By putting together the above statements we obtain the following main result:

Theorem 4.7. *For a small spread $\varepsilon > 0$, and $0 < \mu/\gamma\sigma^2 \neq 1$, the process \tilde{S}_t in Lemma 4.2 is a shadow price. A long-run optimal policy — both for the frictionless market with price \tilde{S}_t and in the market with bid-ask prices $(1 - \varepsilon)S_t, S_t$ — is to keep the risky weight $\tilde{\pi}_t$ (in terms of \tilde{S}_t) in the no-trade region*

$$[\pi_-, \pi_+] = \left[\frac{\mu - \lambda}{\gamma\sigma^2}, \frac{\mu + \lambda}{\gamma\sigma^2} \right].$$

As $\varepsilon \downarrow 0$, its boundaries have the asymptotics

$$\pi_{\pm} = \frac{\mu}{\gamma\sigma^2} \pm \left(\frac{3}{4\gamma} \left(\frac{\mu}{\gamma\sigma^2} \right)^2 \left(1 - \frac{\mu}{\gamma\sigma^2} \right)^2 \right)^{1/3} \varepsilon^{1/3} + O(\varepsilon).$$

The corresponding equivalent safe rate is:

$$r + \beta = r + \frac{\mu^2 - \lambda^2}{\gamma\sigma^2} = r + \frac{\mu^2}{2\gamma\sigma^2} - \frac{\gamma\sigma^2}{2} \left(\frac{3}{4\gamma} \left(\frac{\mu}{\gamma\sigma^2} \right)^2 \left(1 - \frac{\mu}{\gamma\sigma^2} \right)^2 \right)^{2/3} \varepsilon^{2/3} + O(\varepsilon^{4/3}).$$

If $\mu/\gamma\sigma^2 = 1$, then $\tilde{S}_t = S_t$ is a shadow price, and it is optimal to invest all wealth in the risky asset at time $t = 0$, never to trade afterwards. In this case, the equivalent safe rate is the frictionless value $r + \beta = r + \mu^2/2\gamma\sigma^2$.

Proof. First let $0 < \mu/\gamma\sigma^2 \neq 1$. Optimality of the strategy (φ_t^0, φ_t) associated to $\tilde{\pi}(\Upsilon_t)$ for \tilde{S}_t has been shown in Lemma 4.4. The asymptotic expansions are an immediate consequence of their counterpart for λ (cf. Lemma 3.2) and Taylor expansion. Next, Lemma 4.5 shows that \tilde{S}_t is a shadow price process in the sense of Definition 4.1. Proposition 4.6 shows that, for small transaction costs ε , the same policy is also optimal, with the same equivalent safe rate, in the original market with bid-ask prices $(1 - \varepsilon)S_t, S_t$.

Consider now the degenerate case $\mu/\gamma\sigma^2 = 1$. Then the optimal strategy in the frictionless model $\tilde{S}_t = S_t$ transfers all wealth to the risky asset at time $t = 0$, never to trade afterwards ($\varphi_t^0 = 0$ and $\varphi_t = \xi + \xi^0 S_0^0/S_0$ for all $t \geq 0$). Hence it is of finite variation and the number of shares never decreases from the unlevered initial position, and increases only at time $t = 0$, where the shadow price coincides with the ask price. Thus, $\tilde{S}_t = S_t$ is a shadow price. The remaining assertions then follow as in Proposition 4.6 above. \square

The trading boundaries in this paper are optimal for a long investment horizon, but are also approximately optimal for finite horizons. The following theorem, which complements the main result, makes this point precise:

Theorem 4.8. *Fix a time horizon $T > 0$. Then, the finite-horizon equivalent safe rate of the liquidation value $\Xi_T^\phi = \phi_T^0 S_T^0 + \phi_T^+(1 - \lambda)S_T - \phi_T^- S_T$ associated to any strategy (ϕ^0, ϕ) satisfies the upper bound*

$$\frac{1}{T} \log E \left[(\Xi_T^\phi)^{1-\gamma} \right]^{\frac{1}{1-\gamma}} \leq r + \frac{\mu^2 - \lambda^2}{2\gamma\sigma^2} + \frac{1}{T} \log(\phi_{0-}^0 + \phi_{0-} - S_0) + \frac{\mu}{\gamma\sigma^2} \frac{\varepsilon}{T} + O(\varepsilon^{4/3}), \quad (4.16)$$

and the finite-horizon equivalent safe rate of our long-run optimal strategy (φ^0, φ) satisfies the lower bound

$$\frac{1}{T} \log E \left[(\Xi_T^\varphi)^{1-\gamma} \right]^{\frac{1}{1-\gamma}} \geq r + \frac{\mu^2 - \lambda^2}{2\gamma\sigma^2} + \frac{1}{T} \log(\varphi_{0-}^0 + \varphi_{0-} - S_0) - \left(\frac{2\mu}{\gamma\sigma^2} + \frac{\varphi_{0-} - S_0}{\varphi_{0-}^0 + \varphi_{0-} - S_0} \right) \frac{\varepsilon}{T} + O(\varepsilon^{4/3}). \quad (4.17)$$

In particular, for the same unlevered initial position ($\phi_{0-} = \varphi_{0-} \geq 0, \phi_{0-}^0 = \varphi_{0-}^0 \geq 0$), the equivalent safe rates of (ϕ^0, ϕ) and of the optimal policy (φ^0, φ) for horizon T differ by at most

$$\frac{1}{T} \left(\log E \left[(\Xi_T^\phi)^{1-\gamma} \right]^{\frac{1}{1-\gamma}} - \log E \left[(\Xi_T^\varphi)^{1-\gamma} \right]^{\frac{1}{1-\gamma}} \right) \leq \left(\frac{3\mu}{\gamma\sigma^2} + 1 \right) \frac{\varepsilon}{T} + O(\varepsilon^{4/3}). \quad (4.18)$$

This result implies that the horizon, like consumption, only has a second order effect on portfolio choice with transaction costs, because the finite-horizon equivalent safe rate matches, at the leading order $\varepsilon^{2/3}$, the equivalent safe rate of the stationary long-run optimal policy, and recovers, in particular, the first-order asymptotics for the finite-horizon value function obtained by Bichuch (2012, Theorem 4.1).

Proof of Theorem 4.8. Let (ϕ^0, ϕ) be any admissible strategy starting from the initial position $(\varphi_{0-}^0, \varphi_{0-})$. Then as in the proof of Proposition 4.6, we have $\Xi_T^\phi \leq \tilde{X}_T^\phi$ for the corresponding shadow payoff, that is, the terminal value of the wealth process $\tilde{X}_t^\phi = \phi_0^0 + \phi_0 \tilde{S}_0 + \int_0^t \phi_s d\tilde{S}_s$ corresponding to trading ϕ in the frictionless market with price process \tilde{S}_t . Hence, Lemma 5 in Guasoni and Robertson (2012) and the second bound in Lemma 4.3 imply that

$$\frac{1}{(1-\gamma)T} \log E \left[(\Xi_T^\phi)^{1-\gamma} \right] \leq r + \beta + \frac{1}{T} \log(\varphi_{0-}^0 + \varphi_{0-} S_0) + \frac{\gamma}{(1-\gamma)T} \log \hat{E} \left[e^{(\frac{1}{\gamma}-1)(\tilde{q}(\Upsilon_0) - \tilde{q}(\Upsilon_T))} \right]. \quad (4.19)$$

For the strategy (φ^0, φ) from Lemma 4.5, we have $\Xi_T^\varphi \geq (1 - \frac{\varepsilon}{1-\varepsilon} \frac{\mu+\lambda}{\gamma\sigma^2}) \tilde{X}_T^\varphi$ by the proof of Proposition 4.6. Hence the first bound in Lemma 4.3 yields

$$\begin{aligned} \frac{1}{(1-\gamma)T} \log E \left[(\Xi_T^\varphi)^{1-\gamma} \right] &\geq r + \beta + \frac{1}{T} \log(\varphi_{0-}^0 + \varphi_{0-} \tilde{S}_0) + \frac{1}{(1-\gamma)T} \log \hat{E} \left[e^{(1-\gamma)(\tilde{q}(\Upsilon_0) - \tilde{q}(\Upsilon_T))} \right] \\ &\quad + \frac{1}{T} \log \left(1 - \frac{\varepsilon}{1-\varepsilon} \frac{\mu+\lambda}{\gamma\sigma^2} \right). \end{aligned} \quad (4.20)$$

To determine explicit estimates for these bounds, we first analyze the sign of $\tilde{w}(y) = w - \frac{w'}{1-w}$ and hence the monotonicity of $\tilde{q}(y) = \int_0^y \tilde{w}(z) dz$. Whenever $\tilde{w} = 0$, i.e., $w' = w(1-w)$, the derivative of \tilde{w} is

$$\tilde{w}' = w' - \frac{w''(1-w) + w'^2}{(1-w)^2} = \frac{(1-2\gamma)w'w + \frac{2\mu}{\sigma^2}w'}{1-w} - \left(\frac{w'}{1-w} \right)^2 = 2\gamma w \left(\frac{\mu}{\gamma\sigma^2} - w \right),$$

where we have used the ODE (4.6) for the second equality. Since \tilde{w} vanishes at 0 and $\log(u/l)$ by the boundary conditions for w and w' , this shows that the behaviour of \tilde{w} depends on whether the investor's position is leveraged or not. In the absence of leverage, $\mu/\gamma\sigma^2 \in (0, 1)$, \tilde{w} is defined on $[0, \log(u/l)]$. It vanishes at the left boundary 0 and then increases since its derivative is initially positive by the initial condition for w . Once the function w has increased to level $\mu/\gamma\sigma^2$, the derivative of \tilde{w} starts to become negative; as a result, \tilde{w} begins to decrease until it reaches level zero again at $\log(u/l)$. In particular, \tilde{w} is nonnegative for $\mu/\gamma\sigma^2 \in (0, 1)$.

In the leverage case $\mu/\gamma\sigma^2 > 1$, the situation is reversed. Then, \tilde{w} is defined on $[\log(u/l), 0]$ and, by the boundary condition for w at $\log(u/l)$, therefore starts to decrease after starting from zero at $\log(u/l)$. Once w has decreased to level $\mu/\gamma\sigma^2$, \tilde{w} starts increasing until it reaches level zero again at 0. Hence, \tilde{w} is nonpositive for $\mu/\gamma\sigma^2 > 1$.

Now, consider Case 2 of Lemma 3.1; the calculations for the other cases follow along the same lines with minor modifications. Then $\mu/\gamma\sigma^2 \in (0, 1)$ and \tilde{q} is positive and increasing. Hence,

$$\frac{\gamma}{(1-\gamma)T} \log \hat{E} \left[e^{(\frac{1}{\gamma}-1)(\tilde{q}(\Upsilon_0) - \tilde{q}(\Upsilon_T))} \right] \leq \frac{1}{T} \int_0^{\log(u/l)} \tilde{w}(y) dy \quad (4.21)$$

and likewise

$$\frac{1}{(1-\gamma)T} \log \hat{E} \left[e^{(1-\gamma)(\tilde{q}(\Upsilon_0) - \tilde{q}(\Upsilon_T))} \right] \geq -\frac{1}{T} \int_0^{\log(u/l)} \tilde{w}(y) dy. \quad (4.22)$$

Since $\tilde{w}(y) = w(y) - w'(y)/(1-w)$, the boundary condions for w imply

$$\int_0^{\log(u/l)} \tilde{w}(y) dy = \int_0^{\log(u/l)} w(y) dy - \log \left(\frac{\mu - \lambda - \gamma\sigma^2}{\mu + \lambda - \gamma\sigma^2} \right). \quad (4.23)$$

By elementary integration of the explicit formula in Lemma 3.1 and using the boundary conditions from Lemma 3.3 for the evaluation of the result at 0 resp. $\log(u/l)$, the integral of w can also be computed in closed form:

$$\int_0^{\log(u/l)} w(y) dy = \frac{\frac{\mu}{\sigma^2} - \frac{1}{2}}{\gamma - 1} \log \left(\frac{1}{1-\varepsilon} \frac{(\mu+\lambda)(\mu-\lambda-\gamma\sigma^2)}{(\mu-\lambda)(\mu+\lambda-\gamma\sigma^2)} \right) + \frac{1}{2(\gamma-1)} \log \left(\frac{(\mu+\lambda)(\mu+\lambda-\gamma\sigma^2)}{(\mu-\lambda)(\mu-\lambda-\gamma\sigma^2)} \right). \quad (4.24)$$

As $\varepsilon \downarrow 0$, a Taylor expansion and the power series for λ then yield

$$\int_0^{\log(u/l)} \tilde{w}(y) dy = \frac{\mu}{\gamma\sigma^2} \varepsilon + O(\varepsilon^{4/3}).$$

Likewise,

$$\log \left(1 - \frac{\varepsilon}{1-\varepsilon} \frac{\mu-\lambda}{\gamma\sigma^2} \right) = -\frac{\mu}{\gamma\sigma^2} \varepsilon + O(\varepsilon^{4/3}),$$

as well as

$$\log(\varphi_{0-}^0 + \varphi_{0-} \tilde{S}_0) \geq \log(\varphi_{0-}^0 + \varphi_{0-} S_0) - \frac{\varphi_{0-} S_0}{\varphi_{0-}^0 + \varphi_{0-} S_0} \varepsilon + O(\varepsilon^2),$$

and the claimed bounds follow from (4.19) and (4.21) resp. (4.20) and (4.22). \square

5 Open Problems

In this section we mention three problems for which, in our view, the above approach holds promise, and the effect of transaction costs is likely to be substantial. Of course, only future research can shed light on this point.

5.1 Multiple Assets

In sharp contrast to frictionless models, passing from one to several risky assets is far from trivial with transaction costs. The reason is that, since in the free boundary problem the unknown boundary has one dimension less than the number of risky assets, with one asset it reduces to two points only, but with two assets it already becomes an unknown curve. More importantly, multiple assets introduce novel effects, which defy the one-dimensional intuition, as we now argue. For example, consider a market with two risky assets with prices S_t^1 and S_t^2 :

$$\frac{dS_t^1}{S_t^1} = \mu_1 dt + \sigma_1 dW_t^1 \quad (5.1)$$

$$\frac{dS_t^2}{S_t^2} = \mu_2 dt + \varrho \sigma_2 dW_t^1 + \sigma_2 \sqrt{1 - \varrho^2} dW_t^2 \quad (5.2)$$

where $\mu_1, \sigma_1, \mu_2, \sigma_2 > 0$, $\varrho \in [-1, 1]$, and W^1, W^2 are two independent Brownian motions. Even for this simple model with power utility, the solution to the portfolio choice problem is unknown. Some recent papers, e.g., Atkinson and Mokkhavesa (2004), Muthuraman and Kumar (2006), Atkinson and Ingpochai (2006), Law et al. (2007), and Bichuch and Shreve (2011), offer some insights – and raise a number of questions.

Recall that the frictionless portfolio in the above model is $\pi = \frac{1}{\gamma} \Sigma^{-1} \mu$, where $\mu = (\mu_1, \mu_2)$ is the vector of excess returns, and Σ is the covariance matrix defined as $\Sigma_{11} = \sigma_1^2$, $\Sigma_{12} = \Sigma_{21} = \varrho \sigma_1 \sigma_2$, and $\Sigma_{22} = \sigma_2^2$. In other words:

$$\pi_1 = \frac{\mu_1 - \beta_1 \mu_2}{(1 - \varrho^2) \sigma_1^2} \quad \pi_2 = \frac{\mu_2 - \beta_2 \mu_1}{(1 - \varrho^2) \sigma_2^2}$$

where $\beta_i = (\varrho \sigma_1 \sigma_2) / \sigma_i^2$ are the betas of each asset with respect to the other. In particular, for two uncorrelated assets the portfolio separates, in that the optimal weight for each risky asset in the market with all assets equals the optimal weight for the risky asset in a market with that risky asset only. This separation property is intuitive and appealing, and reduces the analysis of frictionless portfolio choice problems with multiple uncorrelated assets to the single asset case. Liu (2004) and Guasoni and Muhle-Karbe (2011) show that such a separation carries over to transaction cost models with exponential utility.

Surprisingly enough, separation seems to fail with constant relative risk aversion, in that the width of the no-trade region for each asset is affected by the presence of the other, even with zero correlation and logarithmic utility. For example, the heuristics in Law et al. (2007) yield the following width for the no-trade region of the first asset, compare their Equation (50):

$$H_1 = \left(\frac{3\varepsilon}{2\sigma_1^2} \left[\left(\frac{1}{2} \mu' \Sigma^{-1} \mu + \sigma_1^2 \right) \pi_1^2 - \mu_1 \pi_1^2 \right] \right)^{1/3}. \quad (5.3)$$

This quantity clearly depends also on μ_2 and σ_2 through the total squared Sharpe ratio $\mu' \Sigma^{-1} \mu$, even with zero correlation, and hence differs from the width of the no-trade region with a single risky asset:

$$h_1 = \left(\frac{3\varepsilon}{4} \pi_1^2 (1 - \pi_1)^2 \right)^{1/3}. \quad (5.4)$$

Further, a simple calculation shows that, if $\varrho = 0$, then:

$$H_1^3 - h_1^3 = \frac{1}{2\sigma_1^4} \left(\frac{\mu_1 \mu_2}{\sigma_1 \sigma_2} \right)^2. \quad (5.5)$$

In other words, the no-trade region in the larger market is always wider than the no-trade region with one asset, and they coincide only if either asset is useless ($\mu_1 = 0$ or $\mu_2 = 0$). In all other cases, the presence of an independent asset increases the no-trade region of the others, presumably because the variation of the position in each asset becomes less important for the overall welfare of the investor than with a single asset. This observation clearly runs against the common wisdom of fund-separation results for frictionless markets, and has potential implications for intermediation and welfare.

Note that in a frictionless market an investor with power utility is indifferent between trading two uncorrelated assets with Sharpe ratios μ_1/σ_1 , μ_2/σ_2 , and a single asset with Sharpe ratio $\sqrt{(\mu_1/\sigma_1)^2 + (\mu_2/\sigma_2)^2}$, that is, squared Sharpe ratios and in turn equivalent safe rates add across independent shocks. The above observation suggests that this property no longer holds with transaction costs, and an important open question is to understand the welfare difference between the

two markets. If the two-asset market is more attractive, then investors benefit from access to individual securities rather than only to a limited number of funds, in contrast to classic fund-separation results. Of course, the question is whether this effect is indeed present and large enough to be relevant.

5.2 Predictability

Can future stock returns be predicted with public information? And what increase in welfare can one expect from this information? Predictably enough, these questions have generated a voluminous literature, which evaluates the statistical significance as well as the in-sample and out-of-sample performance of several predictors that focus either on stock characteristics, such as the dividend-yield and earnings-price ratio, or interest rates, such as the term-spread and the corporate-spread.

Perhaps less predictably, this voluminous literature remains divided between the weak statistical significance of several models, and the strong economic significance of parameter estimates. On the one hand, the standard errors of the predictability parameters are of the same order of magnitude as the parameter estimates themselves; on the other hand, these estimates – if valid – imply a substantial welfare increase. These opposing viewpoints are discussed in Welch and Goyal (2008), who offer a critical view of the empirical literature, and find that most models have poor out-of-sample performance, and Cochrane (2008), who argues that the absence of predictability in dividend growth implies the presence of return predictability.

Kim and Omberg (1996) introduce a basic model with predictable returns, based on one asset with price S_t , and one state variable θ_t :

$$dS_t/S_t = (\mu + \alpha\theta_t)dt + \sigma dW_t, \quad (5.6)$$

$$d\theta_t = -\kappa\theta_t dt + dB_t. \quad (5.7)$$

Here θ_t represents a state variable, like the dividend yield, that helps predict future returns, in that the conditional distribution of S_T/S_t at time t depends on θ_t . The two Brownian motions W and B typically have a substantial negative correlation ρ . The parameter α controls the predictability of returns, with $\alpha = 0$ corresponding to the classical case of IID returns.

In such a market, expected returns change over time, mean-reverting to the average μ . Such variation is detrimental for an investor who adopts the constant policy $\pi = \frac{\mu}{\gamma\sigma^2}$, which is optimal for $\alpha = 0$, because time-varying returns increase the dispersion of the final payoff. However, the investor can benefit from *market timing*, that is the ability to adopt an investment policy that depends on the current value of the state variable θ_t . This point is easily seen for logarithmic utility, for which the optimal portfolio is $\pi_t = (\mu + \alpha\theta_t)/\sigma^2$, and the corresponding equivalent safe rate has the simple formula:

$$\lim_{T \rightarrow \infty} \frac{1}{T} E[\log X_T^\pi] = \frac{\mu^2}{2\sigma^2} + \frac{\alpha^2}{4\kappa\sigma^2}. \quad (5.8)$$

This expression shows that the investor benefits from stronger signals (larger α) and from slower mean reversion of the return rate (smaller κ), and the same conclusion broadly applies to power utility, even though the formulas become clumsier, as the optimal portfolio includes an intertemporal hedging component that is absent in the logarithmic case.

The above calculation underlies most estimates of the economic significance of predictability, but obviously ignores transaction costs. This omission may be especially important, as market timing requires active trading, which in turn entails higher costs. In short, while the potential benefit of predicability is clear from the frictionless theory, its potential costs are blissfully ignored, but may be substantial, and a priori may or may not offset benefits.

Remarkably enough, the above model with transaction costs has never been solved, even for logarithmic utility. Intuitively, the solution of this model should lead to a buy curve $\pi_-(\theta)$ and to a sell curve $\pi_+(\theta)$, which describe the no-trade region for each value of the state variable θ_t . Still at an intuitive level, the width of the no-trade region should be wider for values of θ that are farther from zero, since the portfolio is increasingly likely to return towards the frictionless optimum without trading.

At the technical level, the model includes two state variables: the predictor θ_t , and the current risky weight π_t . The presence of two state variables in turn implies that the value function satisfies an elliptic linear partial differential equation within the no-trade region, along with the boundary and smooth-pasting conditions at the boundary. The difficulty is to characterize the shape of the no-trade interval $[\pi_-(\theta), \pi_+(\theta)]$, as a function of the state θ_t , along with its implied equivalent safe rate.

Solving such a model can contribute to the predictability debate by clarifying the extent to which the ability to *forecast* future returns can translate into the ability to *deliver* higher returns by trading. When transaction costs are included, it may turn out that potential benefits of market timing are minimal, even if return predictability is statistically significant.

5.3 Options Spreads

Options listed on stock exchanges display much wider bid-ask spreads than their underlying assets. While the spread on a large capitalization stock is typically less than ten basis points, even the most liquid at-the-money options have spreads of several *percentage* points. To the best of our knowledge, there seems to be no theoretical work that links the bid-ask spread of an asset to the spread of its options.

Of course, in a frictionless, complete market, both spreads are zero, and the option is replicated by a trading strategy in the underlying asset. The problem is that introducing a bid-ask spread for the underlying asset immediately makes the notion of option price ambiguous. Even if the asset follows a geometric Brownian motion, with transaction costs the superreplication price of any call option equals the stock price itself (Soner et al., 1995). Similarly, the subreplication price is zero. Thus, one cannot interpret the bid and ask prices of the option as replication bounds, if the intention is to obtain a realistic spread.

In contrast to the previous two problems, in which the model is clear and the challenges are mathematical, this question poses some conceptual issues at the outset. One possibility is to interpret the bid and ask prices of the options as marginal prices in a partial equilibrium setting. For example, suppose that the bid and ask prices of the asset are exogenous, and follow geometric Brownian motion, with a constant relative bid-ask spread. Suppose also that a representative investor freely trades this asset, and a European option with maturity T , as to maximize utility from terminal wealth, either at the same maturity, or at some long horizon.

Since options, unlike stocks, exist in zero net supply, assume that the representative investor's optimal policy is to keep a zero position in the option at all times. In a complete frictionless market, this condition uniquely identifies the option price as the unique arbitrage-free price. With transaction costs, it leaves more flexibility in option price dynamics. Indeed, consider the shadow price corresponding to the utility maximization problem. Since the shadow market is complete, the shadow asset price uniquely identifies a shadow price for the option as the conditional expectation under the risk-neutral probability. For an option of European type with payoff $G(S_T)$, the latter will then be a function $g(t, S_t, Y_t)$ of time, the current stock price, and the current value of the state variable measuring the ratio of risky and safe positions.

Now, suppose that to the original (not shadow) market one adds the option, with a price

dynamics equal to the shadow option price, and zero spread. This market is equivalent to the one with the asset only: by contradiction, if some trading strategy delivered a higher utility than the optimum in the asset-only market, the same strategy would also deliver the same or higher utility in the shadow market (by domination), thereby contradicting the definition of a shadow market. Now, the shadow option price depends on the state variable, which is unobservable since market makers cannot see the private positions of market participants. However, taking the pointwise maxima $\bar{g}(t, S_t) = \max_{y \in [0, \log(u/l)]} g(t, S_t, y)$ and minima $\underline{g}(t, S_t) = \min_{y \in [0, \log(u/l)]} g(t, S_t, y)$ over *all* values of $Y_t \in [0, \log(u/l)]$, one can obtain observable upper and lower bounds on the option price, which depend on the asset price alone. Such bounds are natural candidates for bid and ask prices of the option, because they are the minimal observable bounds that an option price needs to satisfy if its net demand has to be zero.

The question is whether this construction can predict bid-ask spreads that are consistent with the ones observed in reality, hence much wider than those of the underlying asset.

References

- C. Atkinson and P. Ingpochai. The influence of correlation on multi-asset portfolio optimization with transaction costs. *J. Comput. Finance*, 10(2):53–96, 2006.
- C. Atkinson and S. Mookhavesa. Multi-asset portfolio optimization with transaction cost. *Appl. Math. Finance*, 11(2):95–123, 2004.
- M. Bichuch. Asymptotic analysis for optimal investment in finite time with transaction costs. *SIAM J. Financial Math.*, 3(1):433–458, 2012.
- M. Bichuch and S.E. Shreve. Utility maximization trading two futures with transaction costs. Preprint, 2011.
- P. Carr. Randomization and the American put. *Rev. Finan. Stud.*, 11(3):597–626, 1998.
- J. Choi, M. Sirbu, and Žitković. Shadow prices and well-posedness in the problem of optimal investment and consumption with transaction costs. Preprint, 2012.
- J.H. Cochrane. The dog that did not bark: A defense of return predictability. *Rev. Finan. Stud.*, 21(4):1533–1575, 2008.
- G.M. Constantinides. Capital market equilibrium with transaction costs. *J. Polit. Economy*, 94(4):842–862, 1986.
- J. Cvitanić and I. Karatzas. Hedging and portfolio optimization under transaction costs: a martingale approach. *Math. Finance*, 6(2):133–165, 1996.
- M. H. A. Davis and A. R. Norman. Portfolio selection with transaction costs. *Math. Oper. Res.*, 15(4):676–713, 1990.
- B. Dumas. Super contact and related optimality conditions. *J. Econom. Dynam. Control*, 15(4):675–685, 1991.
- B. Dumas and E. Luciano. An exact solution to a dynamic portfolio choice problem under transactions costs. *J. Finance*, 46(2):577–595, 1991.
- S. Gerhold, P. Guasoni, J. Muhle-Karbe, and W. Schachermayer. Transaction costs, trading volume, and the liquidity premium. Preprint, 2011a.
- S. Gerhold, J. Muhle-Karbe, and W. Schachermayer. The dual optimizer for the growth-optimal portfolio under transaction costs. *Finance Stoch.*, (To appear), 2011b.
- P. Guasoni and J. Muhle-Karbe. Long horizons, high risk-aversion, and endogenous spreads. Preprint, 2011.
- P. Guasoni and S. Robertson. Portfolios and risk premia for the long run. *Ann. Appl. Probab.*, 22(1):239–284, 2012.
- P. Guasoni, M. Rásonyi, and W. Schachermayer. Consistent price systems and face-lifting pricing under transaction costs. *Ann. Appl. Probab.*, 18(2):491–520, 2008.
- A. Herzegh and V. Prokaj. Shadow price in the power utility case. Preprint, 2011.
- K. Janeček and S. E. Shreve. Asymptotic analysis for optimal investment and consumption with transaction costs. *Finance Stoch.*, 8(2):181–206, 2004.
- E. Jouini and H. Kallal. Martingales and arbitrage in securities markets with transaction costs. *J. Econom. Theory*, 66(1):178–197, 1995.

- J. Kallsen and J. Muhle-Karbe. On using shadow prices in portfolio optimization with transaction costs. *Ann. Appl. Probab.*, 20(4):1341–1358, 2010.
- T.S. Kim and E. Omberg. Dynamic nonmyopic portfolio behavior. *Rev. Finan. Stud.*, 9(1):141–161, 1996.
- D. Kramkov and W. Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Ann. Appl. Probab.*, 9(3):904–950, 1999.
- S.L. Law, C.F. Lee, S. Howison, and J.N. Dewynne. Correlated multi-asset portfolio optimisation with transaction cost. Preprint, 2007.
- H. Liu. Optimal consumption and investment with transaction costs and multiple risky assets. *J. Finance*, 59(1):289–338, 2004.
- H. Liu and M. Loewenstein. Optimal portfolio selection with transaction costs and finite horizons. *Rev. Finan. Stud.*, 15(3):805–835, 2002.
- J. Liu. Portfolio selection in stochastic environments. *Rev. Finan. Stud.*, 20(1):1–39, 2007.
- R. Liu and J. Muhle-Karbe. Portfolio selection with small transaction costs and binding portfolio constraints. Preprint, 2012.
- M. Loewenstein. On optimal portfolio trading strategies for an investor facing transactions costs in a continuous trading market. *J. Math. Econom.*, 33(2):209–228, 2000.
- M. J. P. Magill and G. M. Constantinides. Portfolio selection with transactions costs. *J. Econom. Theory*, 13(2):245–263, 1976.
- R.C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *Rev. Econ. Statist.*, 51(3):247–257, 1969.
- Robert C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *J. Econom. Theory*, 3(4):373–413, 1971.
- K. Muthuraman and S. Kumar. Multidimensional portfolio optimization with proportional transaction costs. *Math. Finance*, 16(2):301–335, 2006.
- L. C. G. Rogers. Why is the effect of proportional transaction costs $O(\delta^{2/3})$? In *Mathematics of Finance*, volume 351 of *Contemp. Math.*, pages 303–308. Amer. Math. Soc., Providence, RI, 2004.
- S. E. Shreve and H. M. Soner. Optimal investment and consumption with transaction costs. *Ann. Appl. Probab.*, 4(3):609–692, 1994.
- A. V. Skorohod. Stochastic equations for diffusion processes with boundaries. II. *Teor. Veroyatnost. i Primenen.*, 7:5–25, 1962.
- H. M. Soner and N. Touzi. Homogenization and asymptotics for small transaction costs. Preprint, 2012.
- H. M. Soner, S. E. Shreve, and J. Cvitanić. There is no nontrivial hedging portfolio for option pricing with transaction costs. *Ann. Appl. Probab.*, 5(2):327–355, 1995.
- M. Taksar, M. J. Klass, and D. Assaf. A diffusion model for optimal portfolio selection in the presence of brokerage fees. *Math. Oper. Res.*, 13(2):277–294, 1988.
- I. Welch and A. Goyal. A comprehensive look at the empirical performance of equity premium prediction. *Rev. Finan. Stud.*, 21(4):1455–1508, 2008.
- A. E. Whalley and P. Wilmott. An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Math. Finance*, 7(3):307–324, 1997.